Separable Least Squares Identification of a Parallel Cascade Model of Human Ankle Stiffness

David T. Westwick[1] and Robert E. Kearney[2]
[1] Dept. Elec. & Comp. Eng., Univ. Calgary, 2500 University Dr. NW, Calgary, AB, T2N 1N4, Canada
[2]Dept. Biomed. Eng. McGill University, 3775 University St, Montréal, PQ, H3A 2B4, Canada

Abstract—**The identification of a dynamic, nonlinear model of human ankle stiffness is considered in a minimum mean squared error framework. The model consists of two parallel pathways, one representing the intrinsic dynamics, the other representing the reflex contribution to the stiffness. The model is shown to be linear in all of its parameters, except for those used to describe a single static nonlinearity in the reflex pathway. A separable least squares optimization algorithm is developed which takes advantage of this structure. This new algorithm is applied to experimental stretch reflex data, and the results compared to the current state-of-the-art algorithm, an iterative technique which fits the two pathways alternately. The relative merits of the two approaches are discussed.**

Keywords— **nonlinear system identification, mean squared optimization, separable least squares, stretch reflex dynamics.**

I. Introduction

System identification methods, which construct mathematical models of dynamic systems from measurements of their inputs and outputs, are often used to study physiological systems. Since these systems are often highly nonlinear, physiological applications often require nonlinear system identification methods [9].

One of the major advances in the field of nonlinear system identification has been the adoption of explicit least squares estimation methods in place of earlier cross-correlation based algorithms [9]. For models, like the Volterra and Wiener series, which are linear in their parameters, the minimum mean squared error (MMSE) solution can be found in closed form, simply by solving a linear regression. However, the number of parameters required to represent high-order kernels limits these methods to relatively low-order systems (second or perhaps third order nonlinearities).

Systems that include high-order nonlinearities are often represented using block-structured models: interconnections of zero-memory (i.e. static) nonlinearities and dynamic linear systems. The simplest of these structures are the Wiener model, a linear dynamic element followed by a static nonlinearity, and the Hammerstein model, a static nonlinearity followed by a linear filter [2].

Block structured models are not very general. Thus, the topology of the model must be appropriate for the system being studied. Figure 1 shows a block structured model that is used to represent the dynamic stiffness of the human ankle [5]. The upper pathway represents the "intrinsic stiffness" of the ankle, the dynamic relationship between its position and the resulting torque, in the absence of any reflexive actions. The lower pathway represents the contribution of the stretch reflex to the overall stiffness. The position input is differentiated, producing the angular velocity of the ankle. This velocity signal is processed by a static nonlinearity, which may represent neural encoding processes in the muscle spindle and in the alpha motoneuron pool. The final element in this pathway is a dynamic linear system, which is thought to represent the contractile dynamics of the muscle.

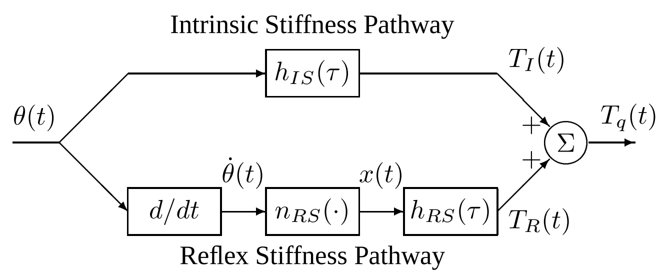

Fig. 1. Parallel Cascade model of the reflex stiffness, the dynamic relationship between the ankle angle, θ, and the ankle torque, T_q. The model includes both intrinsic (upper pathway) and reflex (lower pathway) components.

Like the other block-structured models, this parallel cascade stiffness model (PCSM) isn't linear in its parameters. Thus, unlike the Volterra series, there is no generally applicable closed-form solution for the optimal (MMSE) parameters. However, like the Hammerstein system [10], the PC stiffness model is only nonlinear in the parameters describing its static nonlinearity, and linear in the remaining parameters. Thus, this structure is an ideal candidate for an identification technique based on *separable least squares* optimization.

In this paper, we will develop an identification technique for the PCSM based on a separable least squares optimization. This new technique will be compared to the current state of the art, an iterative, correlation based algorithm, using experimental stretch reflex data. The physiological insights gained from the newly identified models will then be discussed.

II. Theory

The input and output of the PCSM, shown in Fig. 1, are the ankle (angular) position, $\theta(t)$, and the torque, $T_q(t)$, generated about it. The upper pathway represents the intrinsic stiffness. Its output is the torque that would be produced in the absence of reflexes. The transfer function representing dynamic stiffness is improper, in that it has more zeros than poles [4]. In the discrete-time domain, this means that the impulse response of the intrinsic stiffness will be a two-sided filter [1]. Thus,

Report Documentation Page

Report Date	Report Type	Dates Covered (from... to)
25 Oct 2001	N/A	-

Title and Subtitle Separable Least Squares Identification of a Parallel Cascade Model of Human Ankle Stiffness	**Contract Number**
	Grant Number
	Program Element Number
Author(s)	**Project Number**
	Task Number
	Work Unit Number
Performing Organization Name(s) and Address(es) Department Elec & Comp. Engineering University Calgary 2500 University Drive NW Calgary, AB, T2N 1N4 Canada	**Performing Organization Report Number**
Sponsoring/Monitoring Agency Name(s) and Address(es) US Army Research, Development & Standardization Group (UK) PSC 802 Box 15 FPO AE 09499-1500	**Sponsor/Monitor's Acronym(s)**
	Sponsor/Monitor's Report Number(s)

Distribution/Availability Statement
Approved for public release, distribution unlimited

Supplementary Notes
Papers from 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, October 25-28, 2001, held in Istanbul, Turkey. See also ADM001351 for entire conference on cd-rom.

Abstract

Subject Terms

Report Classification unclassified	**Classification of this page** unclassified
Classification of Abstract unclassified	**Limitation of Abstract** UU

Number of Pages
4

the intrinsic stiffness is represented by the two-sided impulse response (IRF).

$$T_I(t) = \sum_{\tau=-L_I}^{L_I} h_{IS}(\tau)\theta(t-\tau) \tag{1}$$

The lower pathway represents the reflex contribution to the total ankle stiffness. The stretch reflex EMG is often modeled as a Hammerstein system [10] with ankle velocity as its input. In the PCSM, the lower pathway consists of a differentiator, which computes the ankle velocity, followed by a Hammerstein system, which represents the stretch reflex.

Let $\dot{\theta}(t)$ be the angular velocity, obtained by numerically differentiating the position, $\theta(t)$, and let the output of the static nonlinearity, $x(t)$, be computed using a degree Q polynomial,

$$x(t) = \sum_{q=0}^{Q} c_q \dot{\theta}^q(t) \tag{2}$$

The linear subsystem in the reflex pathway will be represented by its impulse response.

$$T_R(t) = \sum_{\tau=0}^{L_R} h_{RS}(\tau)x(t-\tau) \tag{3}$$

which, unlike the IRF of the intrinsic stiffness, is causal. The PCSM output is obtained by summing $T_{RS}(t)$ and $T_{IS}(t)$,

$$\begin{aligned} T_q(t) &= \sum_{\tau=-L_I}^{L_I} h_{IS}(\tau)\theta(t-\tau) \\ &\quad + \sum_{\tau=0}^{L_R} h_{RS}(\tau)\left(\sum_{q=0}^{Q} c_q \dot{\theta}^q(t-\tau)\right) \end{aligned} \tag{4}$$

A. Iterative, Correlation-Based Identification

Currently, the elements of the PCSM are identified using an iterative, cross-correlation based scheme. This is possible because of the delays inherent in the system. The two-sided impulse response representing the inherent dynamics dies out well before 40 ms. The reflex component, on the other hand, includes a delay of at least 40 ms, due to propagation delays. Thus, the two contributions can be separated temporally.

The first step in fitting the reflex cascade model is to fit a two-sided IRF, $\hat{h}_{IS}(\tau)$, whose causal dynamics are limited to $\pm 40ms$, between $\theta(t)$, and $T_q(t)$ using a cross-correlation based technique [1]. The output of this pathway, $\hat{T}_I(t)$, is taken to be an estimate of the intrinsic torque.

An iterative, cross-correlation based technique [2] is then used to fit a Hammerstein cascade between $\dot{\theta}(t)$ and the residuals, $T_q(t) - \hat{T}_I(t)$, where $\hat{T}_I(t)$, is the output of the estimated intrinsic dynamics.

Finally, note that the reflex torque will act as noise in the estimation of the intrinsic dynamics. Thus, it may be possible to improve the estimate of the intrinsic dynamics, by fitting a new IRF between $\theta(t)$ and the residuals $T_q(t) - \hat{T}_R(t)$. This iteration continues, alternating between the intrinsic and reflex dynamics, until the model accuracy converges.

B. Separable Least Squares Identification

Parametric optimization methods can also be used to identify the elements of the PCSM. Thus, the objective would be to find the parameter vector,

$$\beta = \left[\begin{array}{ccc} h_{IS}^T & h_{RS}^T & c^T \end{array}\right]^T \tag{5}$$

that minimizes the cost function

$$V_N(\beta) = \frac{1}{2N}\sum_{t=1}^{N}\left(T_q(t) - \hat{T}_q(t,\beta)\right)^2 \tag{6}$$

where $\hat{T}_q(t,\beta)$ is the model output computed using (4), where the values for the IRFs and polynomial coefficients are contained in the parameter vector, β, defined in (5).

In principle, one could use a gradient descent procedure, such as the Levenberg-Marquardt (L-M) algorithm, to solve the minimization. Thus, one would start with an initial estimate, $\beta^{(0)}$, and refine it, in the case of L-M, as follows,

$$\beta^{(k+1)} = \beta^{(k)} + \left(J^T J + \delta_k I\right)^{-1} J^T \epsilon \tag{7}$$

where δ_k is a regularization parameter used to control the convergence rate and stability. The matrix J is the Jacobian, a matrix whose $[t,m]$ entry contains the partial derivative of the model output at time t with respect to the m'th parameter. Thus,

$$J(t,m) = \frac{\partial \hat{T}_q(t,\beta)}{\partial \beta(m)} \tag{8}$$

However, the PCSM has $(2L_I+1)+(L_R+1)+(Q+1)$ parameters, corresponding to the weights in the intrinsic and reflex IRFs and the polynomial coefficients that describe the nonlinearity. Depending on the memory length of the IRFs, and the order of the polynomial, the stiffness model could easily have more than 100 parameters! Thus, it appears that the parametric optimization will have to search over a 100+ dimensional parameter space. Furthermore, each step (7) in the optimization would require the generation and inversion of a 100 by 100 (or larger) matrix. This is clearly impractical.

Note, however, that if the nonlinearity is fixed, the output of the stiffness model (4) is a linear function of the filter weights of the two IRFs. Thus, for a given static nonlinearity, the optimal weights for both IRFs can be found simultaneously in closed form by solving a linear regression.

Subdivide the parameter vector into linear and nonlinear parameters,

$$\beta = \left[\begin{array}{c|c} \beta_l^T & \beta_n^T \end{array}\right]^T = \left[\begin{array}{cc|c} h_{IS}^T & h_{RS}^T & c^T \end{array}\right]^T \tag{9}$$

Then, for any given nonlinear parameters (polynomial coefficients), β_n, the optimal linear parameters (IRF weights), β_l, can be found by solving the linear regression,

$$T_q = X(\beta_n)\beta_l + \epsilon \tag{10}$$

where T_q is a vector whose t'th element is $T_q(t)$, ϵ is a vector containing the errors, and the regression matrix, $X(\beta_n)$, contains both advanced and lagged copies of θ (since the intrinsic

dynamics are two-sided), and lagged copies of the nonlinearity output. Clearly, only the latter columns depend on the choice of polynomial coefficients.

By solving (10), β_l is expressed as a function of the nonlinear parameters. Thus, the model output and cost function can be considered to be functions of the nonlinear parameters only, and written as $\hat{T}_q(t, \beta_n)$ and $V_N(\beta_n)$. As a result, the gradient based optimization need only search for the optimal β_n, reducing the dimension of the search space to $Q+1$. This procedure is known as *separable least squares* optimization [8].

To apply L-M to the SLS problem, the separated Jacobian must be computed. First consider the Jacobian of the unseparated problem, J, which can be computed using the chain rule. Partition it into linear and nonlinear columns, $J = \left[\begin{array}{c|c} J_l & J_n \end{array}\right]$ corresponding to the linear and nonlinear parameters. Then, let P_l be an orthogonal projection onto the column-space of J_l. It can be shown [8] that the Jacobian for the separated problem is given by,

$$J_{sls} = (I - P_l)J_n \tag{11}$$

Thus, to use the L-M optimization method on a SLS problem, the nonlinear parameter vector, β_n, is updated according to (7), but using J_{sls} from (11) as the Jacobian.

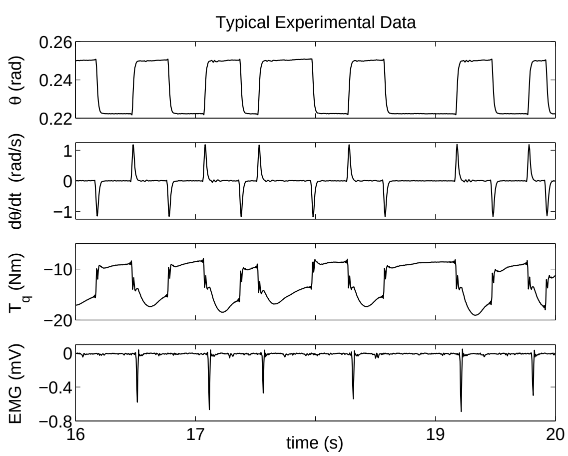

Fig. 2. Extract from a typical experimental trial showing 4 seconds of position, computed velocity, torque and GS EMG.

III. Experimental Results

To evaluate the utility of the SLS algorithm, we used it to estimate intrinsic and reflex stiffness dynamics in a spastic spinal cord injured (SCI) patient. Previously, an iterative, cross-correlation based algorithm has been used to fit the PCSM to experimental data [5], [6]. Tests based on this model are being investigated for possible clinical use [7]. Consequently, methods for efficient, unbiased estimates of its elements are potentially very important.

The experimental paradigm has been described elsewhere [5]. Briefly, the subjects had an incomplete loss of motor function, and clinically evident spasticity due to a previous spinal cord injury[7], which was associated with hyper-active stretch reflexes. They lay supine with their left foot attached to a rotary hydraulic actuator by a custom fitted fiber-glass boot. Ankle torque and position, and the EMG over the Gastrocnemius-Soleus muscles were recorded. The subject maintained a constant background contraction, while a broad-band position perturbation, whose spectrum was shaped to preserve the stretch reflex [5], was applied. Fig. 2 shows 4 seconds of the 30 second (6000 points at 200 Hz) data records used in the analysis.

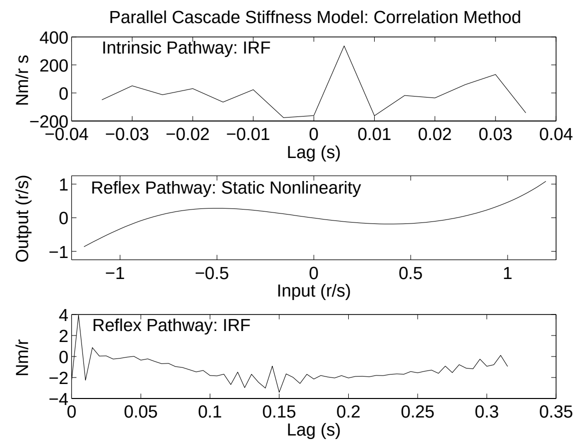

Fig. 3. Elements of the stiffness model identified using the traditional, iterative correlation based algorithm.

We examined the dynamic relationship between the ankle position and torque, by fitting a PCSM, as shown in Fig. 1, between the two signals. Models were fitted between the first 5000 points of the measured position and torque signals. The accuracy of the identified models was then tested by using them to predict the ankle torque measured in the remaining 1000 points, using only the measured position. In the sequel, the initial 5000 points of data will be referred to as the identification segment, whereas the final 1000 points will be called the validation segment.

First, we used the conventional, iterative, correlation based algorithm [5] to fit the intrinsic and reflex pathways of the model. Typical values were chosen for the IRF lengths and polynomial order. Thus, we set the length of the intrinsic dynamics to be ± 35 ms, the length of the reflex dynamics to be 320 ms, and the polynomial order to be 6. Improvement ceased after 3 iterations, producing the model shown in Fig. 3, which fit the identification segment with 83.4 % variance accounted for (% VAF), and predicted the torque in the validation segment with 91.2 % VAF.

Next, we used the separable least squares approach to identify the same model structure. The nonlinearity identified by the iterative correlation based algorithm was used to initialize the SLS procedure. The resulting model is shown in Fig. 4. It has exactly the same structure as the model identified by the iterative correlation based scheme, but accounted for 86 % VAF in the identification segment, and 93.7 % VAF in the validation segment, both significant improvements over the predictions made by the first model.

In comparing the two identified models, the IRFs identified by the conventional scheme include high frequency oscillatory

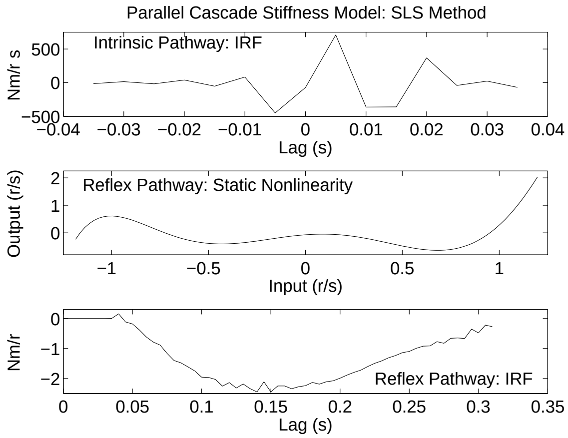

Fig. 4. Elements of the stiffness model identified using the separable least squares optimization based algorithm.

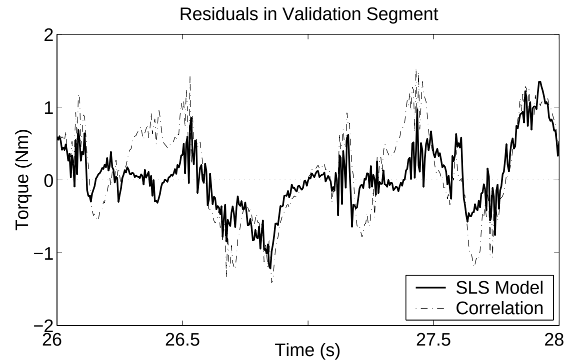

Fig. 5. Errors in the torque predictions during 2 sec. of the validation segment.

components. While these are still evident in the SLS model, they are greatly reduced in amplitude. The nonlinearity identified by the SLS scheme resembles the combination of a small positive threshold, at about 0.8 rad/sec, followed by a half-wave rectifier, an interpretation consistent with empirical studies of the stretch reflex [3]. The functional significance of the nonlinearity identified by the conventional identification scheme is not so evident.

Figure 5 shows the prediction errors due to the two identified models for 2 seconds of the validation segment. The solid line shows the residuals due to the model identified using the SLS technique, whereas the dash-dotted line shows the errors in the output of the model identified using the traditional approach. The improvement in prediction accuracy due to the SLS model is clearly evident in this figure.

IV. Discussion

In this paper, we presented a block structured model for the dynamic stiffness of the human ankle, and developed a SLS identification technique for it. This algorithm was tested on experimental data from a SCI patient, and the resulting model compared to one identified using currently available techniques. The SLS model produced better predictions of the measured torque, both in the training sample, and on validation data. Furthermore, the elements of the SLS model were visibly less noisy, facilitating their physiological interpretation.

This SLS approach is well suited to the parallel cascade stiffness model structure, because the model output is linear in most of its parameters. Thus, using the SLS approach, it is possible to reduce the dimension of the search space dramatically. In this case, the model had 102 parameters (31 weights for $h_{IS}(\tau)$, 64 weights for $h_{RS}(\tau)$, and 7 polynomial coefficients for $n_{RS}(\cdot)$). However, by using the SLS algorithm, it was only necessary to search over a 7 dimensional parameter space, corresponding to the 7 polynomial coefficients.

The biggest disadvantage with using a gradient descent based algorithm, is that the iteration may converge to a sub-optimal local minimum. When using SLS, it is essential to find an initial guess at the nonlinear parameters which is close enough to the globally optimal solution that the search will find it, instead of converging to a sub-optimal solution. In this study, we used the results obtained by the conventional, cross-correlation based method to provide this initial guess.

The static nonlinearity identified by the SLS method included a small positive deflection at relatively large negative velocities. Given the underlying physiology, this is likely to be an artifact, probably due to the use of polynomials in representing the nonlinearity. Other representations, including rational polynomials, splines, or sigmoidal neural networks, may be more appropriate. The incorporation of these nonlinearities into the SLS algorithm is a topic of ongoing research.

Acknowledgements

Supported by the Medical Research Council and Natural Sciences and Engineering Research Council (Canada), and the Univ. Calgary Research Grants Office. M.M. Mirbagheri provided the experimental data.

References

[1] I.W. Hunter and R.E. Kearney. Two-sided linear filter identification. *Med. Biol. Eng. Comput.*, 21:203–209, 1983.

[2] I.W. Hunter and M.J. Korenberg. The identification of nonlinear biological systems: Wiener and Hammerstein cascade models. *Biol. Cybern.*, 55:135–144, 1986.

[3] R.E. Kearney and C.W.Y. Chan. Contrasts between the reflex responses of tibialis anterior and triceps surae to sudden ankle rotation in normal human subjects. *Electroenceph. Clin. Neurophys.*, 54:301–310, 1982.

[4] R.E. Kearney and I.W. Hunter. System identification of human joint dynamics. *CRC Crit. Rev. Biomed. Eng.*, 18:55–87, 1990.

[5] R.E. Kearney, R.B. Stein, and L. Parameswaran. Identification of intrinsic and reflex contributions to human ankle stiffness dynamics. *IEEE Trans. Biomed. Eng.*, 44:493–504, 1997.

[6] M.M. Mirbagheri, H. Barbeau, and R.E. Kearney. Intrinsic and reflex contributions to human ankle stiffness: variation with activation level and position. *Exp. Brain Res.*, 135:423–436, 2000.

[7] M.M. Mirbagheri, R.E. Kearney, and H. Barbeau. Stretch reflex behavior of spastic ankle under passive and active conditions. In *Proc. IEEE EMBS*, 18:2325–2327, 1998.

[8] J. Sjöberg and M. Viberg. Separable non-linear least squares minimization – possible improvements for neural net fitting. In *IEEE Wkshp Neural Nets for Sig. Pro.*, 7:345–354, 1997.

[9] D.T. Westwick and R.E. Kearney. Nonparametric identification of nonlinear biomedical systems, part I: Theory. *Crit. Rev. Biomed. Eng.*, 26:153–226, 1998.

[10] D.T. Westwick and R.E. Kearney. Separable least squares identification of nonlinear Hammerstein models: Application to stretch reflex dynamics. *Ann. Biomed. Eng.*, To Appear, Vol. 29(8).